Arsenii Ashukha

Home page / Google Scholar / GitHub
e-mail: ars.ashuha@gmail.com tel: +79651782915

EDUCATION

- *2017-2021* PhD in Computer Science, **Centre of Deep Learning National Research University HSE**. My PhD is focused on *applications and understanding of stochastic deep learning models*, Advisor: Dmitry Vetrov.

- *2015-2017* MSc in Computer Science, **Moscow Institute of Physics and Technology & Yandex School of Data Analysis**, where I worked with Dmitry Vetrov on learning sparse deep neural networks.

- 2011-2015 BSc in Computer Science **Bauman Moscow State Technical University** with a double major in applied math and computer science. Also, I worked on linguistically motivated topic models with Natalia Loukachevitch.

PROFESSIONAL EXPERIENCE

- **Lead Engineer (eq. Research Scientist)**, Samsung AI Center (2018 - Now):
  The topics I worked on include: a large-scale study of uncertainty estimation techniques and metrics, learning data-dependent priors over models' weights with generative models, learning data augmentation policy for test-time averaging, self-supervised representation learning, transfer learning for RL.
  Since 2021 I'm helping to manage the lab as deputy of lab leader.

- **Researcher**, **Y**andex Research (*2017 - 2018*):
  I worked on a group-level sparsification and uncertainty estimation. The method was applied to accelerate the feature extraction process for image retrieval.

- **Research Intern**, Centre of Deep Learning and Bayesian Methods NRU HSE (2016 - 2017):
  I worked on sparsification of DNNs. This line of research led to Sparse Variational Dropout and started a new class of sparsification methods.

- **Intern**, **Y**andex Music Deep Learning Group (summer of 2016):
  I worked on feature extraction techniques for music data with convolutional neural networks. I also developed an evaluation of learned representations. The representations impacted the content-based recommendation system for music.

- **Machine Learning Engineer**, Rambler&Co (2013 - 2016):
  Worked on demographic classification and recommendation systems. My responsibility included improving the quality and performance of classifiers, automatic feature extraction algorithms, and recommendation algorithms. We used the following stack of technologies: Hadoop, Hive, Spark, XGboost, Vowpal Wabbit, gensim.

REPRESENTATIVE PAPERS

- Arsenii Ashukha*, Alexander Lyzhov*, Dmitry Molchanov*, Dmitry Vetrov
  Pitfalls of In-Domain Uncertainty Estimation and Ensembling in Deep Learning, ICLR (2020)
  blog post / poster video (5mins) / code / arXiv / bibtex

- Dmitry Molchanov*, Arsenii Ashukha*, Dmitry Vetrov
  Variational Dropout Sparsifies Deep Neural Networks, ICML (2017)
  talk / arXiv / bibtex / code theano, tf by GoogleAI, colab pytorch

See the full list at scholar.google.com/citations?user=IU-kuP8AAAAJ, * is for an equal contribution.

CODE
- Check out very short and simple and fan to make implementations of ML algorithms:
  - Gradient Boosting
  - Real NVP
  - Quantile Regression DQN (Distributional RL)
- Also, check out more solid implementations (at least they can do ImageNet):
  - Multi-GPU SimCLRv1
  - Ensembles (Deep ensembles, Snapshot ensembles, cSGLD, FGE, etc.)

REVIEWING
- Conferences:
  - International Conference on Machine Learning, ICML (2019, 2020 top-33% highest-scored reviewers)
  - Neural Information Processing Systems, NeurIPS 2019 (top-50% highest-scored reviewers)
  - International Conference on Learning Representations, ICLR (2020, 2021)
- Workshops:
  - ICML Workshop on Invertible Neural Networks (2019, invertibleworkshop.github.io)
  - Bayesian Deep Learning Workshop (since 2017, bayesiandeeplearning.org)

TEACHING
- Supervisor of reading clubs on machine learning at HSE and Yandex school of data analysis (since 2017)
- Talks and practical sessions at **Deep|Bayes** Summer School on Bayesian Deep Learning (since 2017)
- Lecturer, Moscow Institute of Physics and Technology: I was a lecturer and manager of the deep learning brunch of a facility-wide machine learning course ~60 students (ml-mipt.github.io). Also, I taught deep learning and practical sessions on cutting-edge ML algorithms on a facultative course "Data Mining in Action" ~ 200 students (https://bit.ly/3eRLGYp). The goal of this course is to make ML education available for everyone for free.

SUPERVISION:
- Alexander Lyzhov (moved to Samsung AI Center), Deep Neural Network Ensembles: Analysis and Approaches to Diversification (MSc, 2020)
- Andrei Atanov (PhD candidate at EPFL), Effective Learning of Deep Neural Networks Ensembles (BSc, 2018), Learning Deep Models with Small Data (MSc, 2020)
- Evgenii Nikishin (PhD candidate at Cornell), Stability Improvement and Knowledge Transfer in Deep Reinforcement Learning (MSc, 2019)

FRAMEWORKS & PROGRAMMING LANGUAGES:

- I'm fluent with Python which is my love, I use to code on C, Go, language is not a problem after all.
- I'm also fluent with common python libs such as NumPy, Matplotlib, scikit-learn, etc.
- My primary deep learning framework at the moment is PyTorch, which is my absolute love ❤️ mainly for its simplicity. I love to keep things as simple as possible. Prior to that, I had a decent experience with Theano+Lagange and TensorFlow.
- I have experience with MapReduce, Hadoop, Hive, and Spark.